Survey

Interim Report

on a

Sociolinguistic Survey of Kohistan

At the southern edge of the mountain mass of northern Pakistan lies a region popularly known as Kohistan. This area consists largely of high valleys in the upper reaches of the Panjkora and Swat river systems and along the middle Indus. This is the home of several ethnic groups (in addition to the more recently arrived Gujurs) who speak languages called Kohistani, part of the Dardic branch of the Indo-European family. This branch of Indo-European is said to form a linguistic link between the Indic languages, which are located largely to the east, and the Iranian languages, located largely to the west.

It is in the upper Swat Valley that members of the Sociolinguistic Survey Project sponsored by the Summer Institute of Linguistics has begun a survey of the languages of the northern part of the nation. During the latter half of 1986 a team of three foreign and three Pakistani researchers spent periods of varying length in communities where the Torwali Kohistani, Kalami Kohistani, and Gujuri languages are spoken. After an initial period of language learning and of observing patterns of language use, the team interviewed members of the communities concerning their language use practices and preferences. They also conducted several tests to evaluate the extent to which speakers of the local languages understand both other local languages and the languages of wider communication.

1. Local languages. Most of the people living in the villages of this region speak vernacular languages, most of those unwritten. These languages are in most cases related, but in the survey the relationships among these local languages were studied from two points of view: (a) similarity of lexicon and (b) degree of intelligibility between dialects.

A standard list of 220 words was collected from several areas: two communities in the Torwali area, two communities in the Kalam area in the upper end of the Swat Valley, three communities across the pass in Dir Kohistan, a community of settled Gujurs along the Swat Valley highway, a summer camp of migrant Gujurs, and two Gujur communities outside of Kohistan.

Many lexical items in regular use in these languages have been borrowed from other languages of the region, especially from Pashto and Urdu. From a comparison of the items of the word lists which appear not to be loans it is clear that the following groups of dialects are lexically very similar: (a) (Torwali) Kohistani of Behrain and Chail (95% similar), (b) (Kalami) Kohistani of Kalam, Ushu, Thall and Lamuti (at least 90% similar), (c) Gujuri of Peshmal (Swat) and Balakot (Hazara) (90% similar). On the other hand, these comparisons show that the Kalami Kohistani dialects are quite different lexically from Torwali Kohistani (56%-67%), and that both the Torwali and Kalami groups of dialects are different from Kalkoti (62%-76%). (A preliminary study of the vocabulary of Kohistani dialects in the Indus Valley shows that they are also distinct from the Kalami, Torwali, and Kalkoti dialect groups with 49%-59% similarity.) The dialect of the Ajiri (itinerant) Gujurs is rather different from that of the Gujurs of Balakot (79%). (The similarity of the Ajiri and Peshmal dialects [93%] may be a result of convergence based on frequent contact.) The Gujuri dialect of Gujranwala is even more distinct (only 44%-53% similar). The lexical similarities and differences are summarized in the following listing:

1. Torwali Kohistani: Behrain, Chail
2. Kalami Kohistani: Kalam, Ushu, Thall, Lamuti
3. Kalkoti Kohistani
4. West bank Indus Kohistani: Khandia, Duber, Pattan, Seo
5. Gujuri of Peshmal and Balakot
6. Ajiri Gujuri
7. Gujranwala Gujuri

The gathering of further lexical materials from Kalkot, Indus Kohistan, and perhaps other Gujuri-speaking communities can be expected to further refine this picture of lexical similarity and difference.

The relationships among local languages were also studied from the standpoint of intelligibility. That, it was of interest to know the extent to which speakers of one vernacular language understand speakers of other vernacular languages.

A personal-experience narrative was tape recorded in each dialect to be studied, and the recorded text was played for other speakers of the same dialect to be sure that they could perfectly understand the narrative and verify that the text was a representative sample of their

phase of the study differs from the study of local dialects in several respects. One such difference relates to the sample of the society to be studied. Where intelligibility among related dialects is concerned, (apart from those who have had more than normal interdialectal contact and have learned the second speech forms) it can be safely assumed that one adult speaker of dialect B understands dialect A about as well as another. When speakers of a local language have learned a second language that is unrelated or only remotely related, it is to be expected that individuals will have learned the second language to an unequal degree. Therefore, the selection of the representative sample to be tested is a much more delicate issue than in the case of straight dialect-intelligibility testing.

In order to deal with this question in the community of Kalam, a sociolinguistic profile was developed: Town leaders were consulted, and representatives of over 40 households from different parts of the community were interviewed. In that way it was possible to learn the composition of each household. Information was gathered about each member of the household such as age, sex, place of birth, mother tongue, level of education, occupation, and previous places of residence. From this information it was possible to develop a profile of the community, estimating the portion of the population included in each sociolinguistic sub-group. One such sociolinguistic sub-group is educated young men. They are apt to be more willing and interested in dealing with researchers from outside the community. Normally, they are also among the most proficient second-language speakers in the community. If the sample for bilingualism studies is drawn largely from this sub-group, the conclusions drawn for the entire community are likely to be seriously skewed. In Kalam, educated (at least 5 years of school), young (21-30 years old) men did exhibit a high proficiency in Urdu. However, this fact is seen in proper perspective when it is realized that this visible and influential segment constitutes only about 2% of the community.

Two types of instruments were used in evaluating levels of proficiency in Urdu and Pashto. Texts recorded by mother-tongue speakers of Urdu or Pashto were developed as recorded-text tests, similar to those described above for dialect-intelligibility testing. A second type of test involved a series of sentences in Urdu or Pashto of varying difficulty, which each second-language speaker tested was asked to repeat. His repetitions were scored for accuracy, and his total score was taken as an index of his proficiency in hearing and speaking that second language.

speech. Later, the recording was played for speakers of other local dialects or languages. A group of questions about the content of the story, which had been previously translated into the dialect of the people being tested, was used to test intelligibility of that story. After they had listened to the recorded story, the subjects listened to each question in their own dialect and answered those questions according to their understanding of the recorded text. If a subject understood the story in the other dialect perfectly, he was able to answer all questions perfectly. A more limited understanding was reflected in his answering some questions correctly and some incorrectly or not at all. After about ten people in a community had listened to the story and answered the questions, a percentage of correct answers was calculated. This figure can be taken as an index of the extent to which speakers of dialect B understand dialect A.

Since a high degree of lexical similarity was found between Peshmal and Ajiri Gujuri, it is hardly surprising that those Ajiri speakers who listened to the Peshmal Gujuri text were able to answer all questions correctly, indicating that Ajiri Gujurs understand Peshmal Gujuri very well.

There was found to be a much greater problem for some speakers of Kohistani in the Swat Valley to understand other speakers of Kohistani. The Kohistani speakers of Kalam understood an average of only 60% of the text of the Torwali Kohistani of Behrain.

It often occurs that speakers of one dialect (B) do not understand another dialect (A) at the same level that speakers of A understand B. And so it was in Swat Kohistan, where Kohistani speakers from Behrain understood an average of only 43% of the Kohistani text from Kalam, whereas speakers from Kalam understood 60% of the Kohistani text from Behrain. In both cases, however, the scores were so low that it is clear that intelligibility between these dialects is very poor. This conclusion is further confirmed by the fact that Kohistanis from Kalam report that they require contact with Torwali Kohistani over a period of several years before they can understand it well. It is also confirmed by the fact that Kohistanis from Kalam speak Kohistani with people from Thall and Lamuti but speak Pashto with speakers of Torwali Kohistani.

2. Use of regional and national languages. A second major focus of the sociolinguistic survey deals with the proficiency with which speakers of Kohistani or Gujuri control Pashto and Urdu. This

In Kalam the educated men performed better on both kinds of Urdu tests than did the uneducated men. There was only a small variation in the scores of the educated men, probably reflecting the fact that most of their learning has taken place in the same context, namely school. The uneducated men performed less well. The sub-group of uneducated men with the highest average score (83%) on the recorded-text test in Urdu had scores lower than the sub-group of educated men with the lowest average score (88%). Also, the range of scores among uneducated men was considerably greater, reflecting the variety of contexts in which they have learned Urdu and the differing degrees of opportunity they have had.

It should be noted that while the performance of the uneducated men on the recorded-text test was definitely lower than that of the educated men, on the sentence-repetition test their performance was lower by an even greater margin. This suggests that the sentence-repetition test, which requires ability to speak as well as to understand, is a more sensitive indicator of bilingual proficiency at middle and lower levels. In the developing of sociolinguistic testing instruments, of course, it is of considerable interest to have at one's disposal techniques which are capable of discriminating at various ranges of the scale of bilingual proficiency.

It is of further interest that the text tests in Urdu showed some difference in scores between educated (89.5%) and uneducated (73.5%) men in the Kalami and Torwali areas, whereas much less difference was found in the Pashto test scores between educated (95.7%) and uneducated (91.6%) men. This highlights the fact that Pashto, although used as a medium of instruction in the schools, is not primarily a language learned in school, whereas proficiency in Urdu is clearly related to level of schooling. This point is further underscored by the fact that the uneducated men in Kalam actually scored slightly more than one percentage point higher in the Pashto-text test than did their educated fellows.

As a result of comparing scores on bilingual proficiency tests with the sociolinguistic factors of the test subjects, three factors have emerged thus far as being significant in promoting proficiency in a second language: age, level of education, and travel to lower areas in winter.

Without exception the middle-aged men (31-45 years) have the highest scores in recorded text tests in Urdu and Pashto. In some cases their scores are slightly higher than the two younger groups of men, in other cases dramatically higher. In every case the overall scores of the

oldest men (46 years or older) are lower than those of the middle-aged men. However, when we distinguish educated men from uneducated, we find without exception that the scores of the educated older men remained high. In fact, in every case they had perfect scores. By contrast, the scores of the older uneducated men dropped significantly. This may reflect the fact that middle-aged men have more need and/or resources for travel to areas where they must use Urdu or Pashto on a prolonged basis. The proficiency scores suggest that inter-ethnic contacts continue for the educated men into older age but that they drop off significantly for uneducated men. This, however, is yet to be verified.

The effect of education on the performance of men in the tests, especially in Urdu, has been repeatedly noted above. However, it is significant that this potent factor of education applies to only about 7% of the population. Thus, it is a highly restricted factor.

A much more common factor in promoting bilingual proficiency is travel to lower elevations during winter, when the higher elevations are often snowbound. During the course of conducting interviews in Kalam only two men reported that they had not traveled away from the area. Most of them sometimes travel to the Lower Swat Valley or Peshawar, which favors Pashto bilingualism, and sometimes to the Punjab, which favors Urdu.

It is clear that in the communities of Upper Swat Valley the local language predominates as the medium of intra-ethnic communication, whereas Pashto predominates in inter-ethnic communication. Eighty-five percent of those interviewed in Kalam use only Kalami Kohistani in the home. The other fifteen percent report using Kalami Kohistani in addition to Pashto or Torwali Kohistani. One man stated that only about 25 Kohistani women in the community of about 12,000 can speak Pashto well.

Interviews in Peshmal revealed that many parents have no objection to their children's marrying Pashto-speaking persons. However, some object to their sons' marrying Pashto-speaking women because when they come to live in their husbands' homes they cannot communicate. They explained that Gujur women do not speak Pashto. All of this points to a vigorous use of the local languages in the Upper Swat Valley.

Regarding communication with people from outside their areas, eighty-five percent of the respondents from Kalam reported speaking Pashto to Kohistanis from the Indus Valley. Similarly, the Gujurs of Peshmal use Pashto in communicating with Pathans and neighboring Kohistanis. There are, however, two exceptions to this picture of

predominance of Pashto as the language of inter-ethnic contact. Both Kohistanis and Gujurs travel to the Punjab at times either to sell produce or to work during the winter season. There they speak Urdu. Also, a surprising 29% of those interviewed in Kalam reported using Torwali when speaking with Kohistanis from the Behrain area, and an additional 12% reported using either Torwali or Pashto in such contacts.

Attitudes toward the local languages appear to be almost entirely positive. All Kohistani respondents in Kalam stated that they wanted their children to learn Kohistani and that they would like to have their language used in written form. All Gujur respondents in Peshmal stated that they are proud to be heard by others speaking Gujuri in public and that they are never embarrassed by their language. At the same time they reflect positive attitudes toward other languages in the area, especially toward Urdu. When people of Kalam were asked which language they want to know better, 43% said they want to know Urdu better, 48% said English, and 15% said Arabic (allowing for more than one response in some cases). When they were asked which language they want their children to learn, 55% responded that they want them to learn Urdu, 45% answered Pashto, 40% answered English, and 15% answered Arabic. When asked which language they like to listen to most, Urdu was by far the most frequent response (65%, again allowing for multiple responses).

3. Topics yet to be studied. At several points it has been clear that statements regarding levels of bilingual proficiency and to some extent patterns of language use reflect the situation only with men. Because all the field researchers have been men, there has been no opportunity to interview women. Now that two women have joined the survey field-research team, we look forward to conducting studies that will supplement the current picture with data concerning the use of second languages by women.

In December 1986 a first field trip was made to Indus Kohistan. Lexical comparison indicates that the Kohistani dialects spoken in the Indus Valley are quite distinct from those of Swat and Dir Kohistan. However, considerable study remains before a clear picture will be available concerning patterns of relationship among the Indus dialects and between them and the other Kohistani dialects to the west and concerning the extent to which Indus Kohistani people also use regional and national languages.

At least two communities will require further study in the Swat and Panjkora Valleys before the vernacular-language picture for those areas will be clear. Lexical comparisons indicate that Kalkot in Dir Kohistan is distinct from the Kohistani of the Kalam and Thall areas. The people of Kalkot are reported to understand those of Kalam but not vice versa. This suggests low inherent intelligibility with learning of the Kalami dialect on the part of the people of Kalkot. Intelligibility testing needs to be carried out to verify this assumption.

In the community of Ushojo in the Chail side valley near Behrain a distinct dialect is reported. Data are needed to confirm whether this dialect is part of the Kohistani language network or whether it is a part of a language group farther afield.

Once the picture for Kohistani and Gujuri has been rounded out, we plan to turn our attention to surveying the situation of language relationships and proficiency in language use in the Hindko-, Pahari-, and Shina-speaking areas--and eventually throughout northern Pakistan.

Calvin R. Rensch

Project Director